UNDERSTANDING BIOSYNTHESIS OF MAJOR CHEMICAL CONSTITUENTS IN *SWERTIA CHIRAYITA*

SYNOPSIS

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy
In
Biotechnology

By

JIBESH KUMAR PADHAN
Enrolment No. 116559

Under the supervision of

Supervisor: Prof. (Dr.) Rajinder Singh Chauhan

Co-supervisor: Dr. Hemant Sood

DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT-173234, HP, INDIA
FEBRUARY 2017

UNDERSTANDING BIOSYNTHESIS OF MAJOR CHEMICAL CONSTITUENTS IN *SWERTIA CHIRAYITA*

JIBESH KUMAR PADHAN

Introduction

Swertia chirayita (Roxb. ex Fleming) H. Karst., commonly known as Chirata (Family: Gentianaceae), is a biennial/pluri-annual herb of temperate Himalayas found at an altitude of 1200–3000 m (4000–10,000 ft.). The extract of this valued herb is reported to possess antimalarial, antihepatotoxic and hypoglycemic, anticarcinogenic, antiviral, antioxidative, antimicrobial and anti-inflammatory properties. Commercially available herbal formulations such as AYUSH-64, Diabecon, Citazil, Lekoril, Lekorin, Livnol, Malarin, etc., contain Chirata extracts for the treatment of a large number of disorders. **Fig. 1** shows a mature plant of *Swertia chirayita*.

Figure 1. A mature *Swertia chirayita* plant

The broad-spectrum biological activities of Chirata are attributed to the presence of phytochemicals belonging to diverse classes such as xanthones and their derivatives, flavonoids, terpenoids, iridoidsand secoiridoid glycosides [1]. The major phytochemicals of the bitter-tasting plant include swertiamarin, amarogentin (secoiridoid glucosides) and mangiferin, a xanthone *C*-glucoside [2]. Swertiamarin is reported to be effective against hepatitis and shown to exhibit anti-diabetic, anticancer and anti-arthritic activities. Amarogentin is known to be anti-diabetic, anticancer and antileishmanial, whereas mangiferin has been tested for its anti-diabetic, antiatherosclerotic, anticancer, anti-HIV and antiparkinsonian behaviors. Overexploitation for usage in herbal drugs along with other factors such as low seed viability and narrow geographic occurrence has led to the endangered status of the plant.

Despite its huge potential in the herbal drug market very few attempts have been made at molecular level to understand the biosynthesis pathways of major phytochemicals in *Swertia*. Swertiamarin and amarogentin take the classical MVA/MEP route of terpene biosynthesis followed by the secoiridoid pathway (**Fig. 2**). These secoiridoids originate from a common intermediate sweroside, which is biosynthesized from geranyl diphosphate (GPP), a condensation product of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Then they are transformed to the end products through a cascade of chemical conversions along with the formation of intermediates such as irido-dial/trial, deoxyloganic acid, loganic acid and secologanic acid [3, 4]. Swertiamarin is the hydroxylation product of sweroside, whereas amarogentin is the biphenylcarboxylic acid derivative of the same. The biphenylcarboxylic acid moiety was presumed to be biosynthesized from *m*-hydroxybenzoyl-CoA by a polyketide-type pathway from phenylalanine via cinnamic acid and benzoic acid as reported in *Swertia japonica* [5]. A retrobiosynthetic ^{13}C NMR study on *S. chirayita* also proved the formation of biphenylcarboxylic acid moiety from *m*-hydroxybenzoic acid, the latter being formed preferentially from an early shikimate pathway intermediate, rather than from phenylalanine via cinnamic acid and benzoic acid [4]. However, minor fractions of phenylalanine-derived metabolites might also contribute to amarogentin biosynthesis, which awaits an experimental proof.

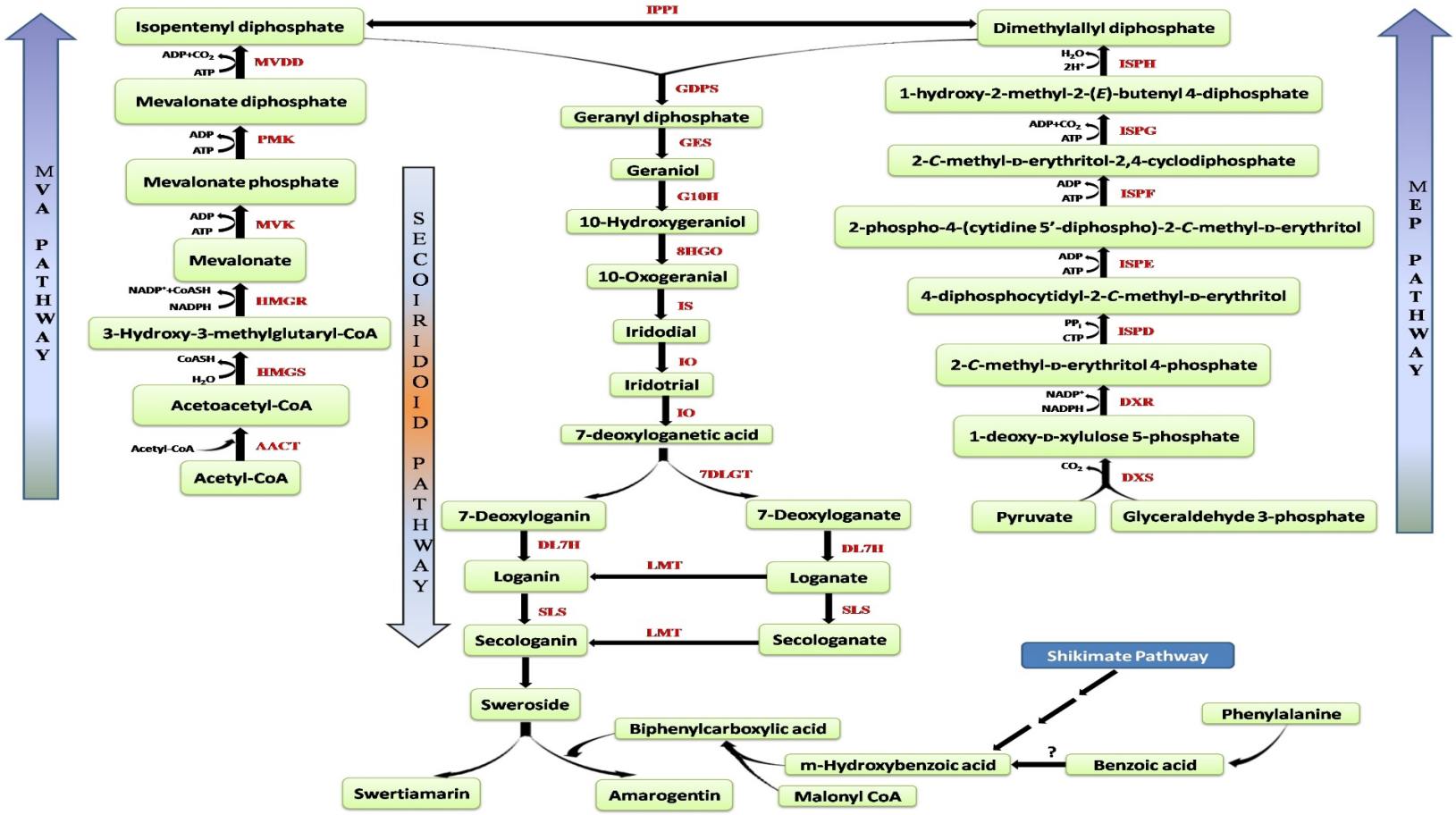

Figure 2. Schematic representation of the proposed swertiamarin–amarogentin biosynthesis pathway (partly adapted from [3]).

Mangiferin, on the other hand, follows a combined phenylpropanoid/ acetate route (**Fig. 3**), where it is biosynthesized from phenylalanine through *p*-coumaric acid/cinamic acid and malonic acid along with benzophenone intermediates such as iriflophenone and maclurin, as described in a study on *Anemarrhena asphodeloides* [6].

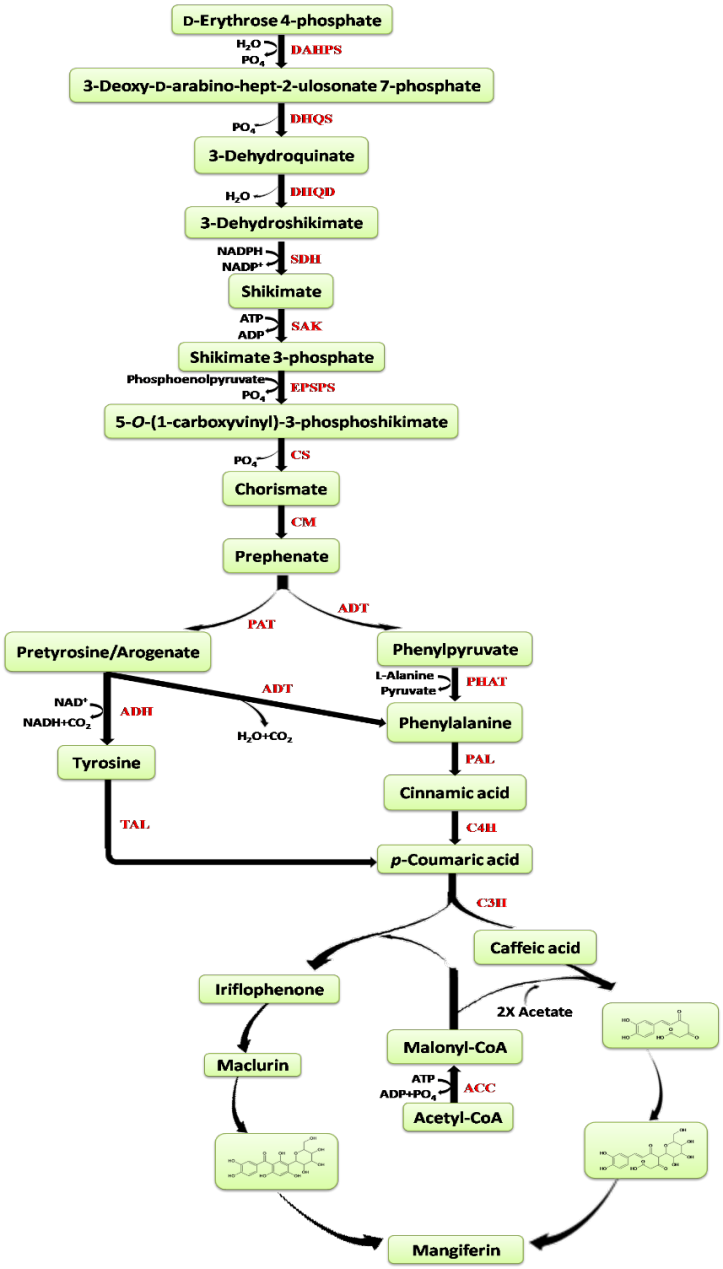

Figure 3. Schematic representation of the proposed mangiferin biosynthesis pathway.

Despite these reports, a complete picture of the molecular understanding of biosynthesis routes leading to the production of major chemical constituents is still lacking. Also, no comprehensive study has been done to associate the expression status of swertiamarin, amarogentin and mangiferin pathway genes with their biosynthesis in *S. chirayita*. Expression analysis of pathway genes in relation to the metabolite levels would give a clue to the relative contribution of each gene on the biosynthetic machinery of these major metabolites. Key genes identified in these studies along with the understanding of various transcription regulators such as miRNA and transcription factors would be helpful in devising molecular strategies for genetic improvement of *S. chirayita* with enhanced secondary metabolite production. Significant variation in the metabolite levels of naturally grown plants and *in vitro* cultured plants have been observed [7]. Comparative analysis of various processes that contribute to the final levels of phytochemicals would not only identify molecular traits altered in these conditions but also provide adequate resources towards formulating better *in vitro* strategies.

Keeping in view the pharmacological importance of *Swertia chirayita*, obscure knowledge on the biosynthesis pathways and the regulators thereof, and the lesser known molecular processes channeling towards secondary metabolite levels, the current research work was carried out with the following objectives;

Objective 1: Assessment of the relative contribution of biosynthesis pathway genes *vis-à-vis* major secondary metabolites in *S. chirayita*

Objective 2: Discerning the effect of altered physiological and biochemical processes in *in vitro* versus *in vivo* grown plants on production levels of secondary metabolites in *S. chirayita*

Objective 3: Investigating the role of miRNA mediated regulatory surveillance on secondary metabolites biosynthesis in *S. chirayita*

Expression analysis of biosynthesis pathway genes in relation to major secondary metabolites contents was carried out in five different tissues of *S. chirayita* varying in contents of major phytochemicals to identify key genes of the pathways. To ascertain the role of various altered physiological and biochemical processes related to secondary metabolites biosynthesis

and the molecular components thereof, comparative transcriptome analysis was done followed by expression studies of pathway genes. Finally, transcriptomes were mined for microRNAs and their possible associations in secondary metabolites biosynthesis were assessed.

Materials and methods

Plant material and tissue collection

Plants of *S. chirayita* were maintained in a greenhouse at the Jaypee University of Information technology, Waknaghat, H.P., India (1500 m altitude; 31°01'N, 77°04'E) with controlled environmental conditions [light (intensity 300–2000 lx), temperature (25 ± 2 °C), relative humidity (≈75%) and photoperiod (14 h day/10 h night)]. Tissue cultured *S. chirayita* plants were maintained at 25 ± 2 °C in media supplemented with Murashige and Skoog (MS) [8] salts, growth hormones [Kinetin (2 mg/l), Indole-3-butyric acid (2 mg/l) and gibberellic acid (3 mg/l)], sucrose (30 g/l), agar–agar (8.5 g/l), pH 5.6–5.7 and controlled environmental conditions [light (intensity 2000 lx), temperature (25 ± 2 °C), relative humidity (≈75%) and photoperiod (16 h day/8 h night)]. Tissue samples from 2-year-old greenhouse plants (roots, leaves and flowers) and 1.5-month old tissue cultured plants (roots and leaves) were collected for expression analysis of biosynthesis pathway genes. Again, for comparative transcriptome studies tissue samples (roots and shoots) with contrasting metabolite levels were procured from plants raised in greenhouse and *in vitro* conditions. All the tissues were collected, immediately frozen in liquid nitrogen and stored at -80 °C until further use.

Extraction and estimation of swertiamarin, mangiferin and amarogentin

Swertiamarin, mangiferin and amarogentin contents were estimated simultaneously by reverse phase High Performance Liquid Chromatography (HPLC Waters 515) as described previously [9]. The plant materials were ground in liquid nitrogen and suspended in 80% methanol. The sample mixtures were sonicated for 10 min. The following day, the samples were centrifuged at 10,000 rpm for 15 min. The supernatant was then filtered through 0.22 µm filter. The desired separation and baseline was obtained in a gradient method, where mobile phase A was composed of 0.1% TFA in water and mobile phase B was a mixture of acetonitrile/water in a ratio of 70:30. The linear gradient at a flow rate of 1.0 ml/min. was started with 15% B; 20% B in next 5 min., 70% B in next 25 min., hold for 5 min.; 15% B in next 5 min., equilibrated for 5

min. and detected at 240 nm UV wavelength. The separation was obtained on Waters Spherisorb 5 µm ODS2 (250 mm 4.6 mm). The compounds were identified on the basis of their retention time and comparison of UV spectra with the authentic standards procured from ChromaDex (Irvine, CA). The quantification was done in triplicate and the data were interpreted in terms of % of fresh weight (fr. wt) of swertiamarin, mangiferin and amarogentin.

Computational identification of miRNAs and their targets in *Swertia* transcriptomes

To search for potential miRNAs in *Swertia* transcriptomes, mature plant miRNAs and their stem-loop structures were obtained from plant microRNA database (PMRD). The foldback miRNA sequences were used for predicting homologous precursor miRNAs in *S. chirayita* transcriptomes using BLAST-2.2.14, obtained from NCBI Genbank at E-value threshold of $1e^{-5}$. The corresponding mature miRNAs were then used for mining possible targets using plant small RNA target analysis server, psRNATarget (http://plantgrn.noble.org/psRNATarget/). Criterion described by Das and Mondal [10] and Xie et al. [11] were followed in the target prediction process. To investigate the possible role of the miRNAs, the targets were aligned to KEGG databases using default settings. Possible targets with their link to primary and secondary metabolism were chosen for further studies. Finally, the precursor miRNAs associated with the filtered target set were subjected to hairpin structure prediction using the Zuker folding algorithm with MFOLD 3.1. Only those miRNAs and the targets thereof, were chosen for expression studies which satisfy the following criteria; (1) the sequence could fold into an appropriate stem-loop hairpin secondary structure; (2) the small RNA sits in one arm of the hairpin structure; (3) no more than 6 mismatches are between the predicted mature miRNA sequence and its opposite miRNA (miRNA*) sequence; (4) no loop or break in miRNA sequences, and (5) predicted secondary structure has higher MFEI and negative MFE. The final set of miRNAs and their targets were grouped according to their common- and uniqueness across the two transcriptomes and finally their expression status was assessed using reverse transcription quantitative PCR (RT-qPCR).

Isolation of total RNA and cDNA synthesis

Total RNA was extracted from the aforementioned tissues using Trizol Reagent (Ambion), according to the manufacturer's instructions. Quality, purity and concentration of RNA were estimated spectrophotometrically by A260 and A280 measurements (NanoDrop, Thermo Scientific) followed by visualizing on ethidium bromide stained agarose gel. Total RNA (1 μg) was used to synthesize first strand cDNA using Verso cDNA synthesis kit with anchored oligo(dT) primers and Verso RT Enhancer (Thermo Scientific, Wilmington, DE) to remove residual DNA contamination, following the manufacturer's protocol. For miRNA expression studies cDNA was prepared using NCode^{TM} miRNA First-Strand cDNA Synthesis (Life technologies) by selective poly-A tailing of miRNA and subsequent cDNA synthesis from poly-A miRNAs.

Expression analysis using RT-qPCR

To correlate the expression status of biosynthesis pathway genes with the secondary metabolite levels, tissue samples from 2-year-old greenhouse plants (roots, leaves and flowers) and 1.5-month old *in vitro* cultured plants (roots and leaves) with varying levels of metabolites were selected. Gene sequences of the MVA/MEP, secoiridoid and phenylpropanoid/acetate pathways, fundamental processes (glycolysis, TCA cycle, oxidative pentose phosphate pathway and Calvin cycle of photosynthesis) and miRNA targets were retrieved from the whole transcriptome data of *S. chirayita* using BLAST similarity search program (http://www.ncbi.nlm.nih.gov/BLAST). Gene specific primers were designed from the retrieved gene sequences using Primer3 [12]. Multiple internal control genes were used to normalize the expression data. A set of six candidate reference genes encoding actin, eEF1a, β-tubulin, histone H3, 26S rRNA and GAPDH were taken into consideration for their expression across different tissues under study and their suitability as reference genes were analyzed using BestKeeper [13]. For normalization of miRNA expression, 5.8S rRNA was used as an internal control. RT-qPCR reactions were run on a CFX96 system (Bio-Rad Laboratories, Hercules CA) with the iScript one step RT PCR kit (Bio-Rad). The cycling conditions for expression analysis of pathway genes, fundamental processes and miRNA targets comprised denaturation for 5 min. at 94 °C, followed by 40 cycles each of denaturation for 20 s at 94 °C, annealing for 30 s at different temperatures, followed by elongation step for 20 s at 72 °C. miRNA expression was carried out using SYBR®

GreenER™ qPCR SuperMix Universal (Thermo Fisher Scientific) following the manufacturer's protocol. All reactions were run in triplicate. Relative fold changes were determined from the *C*q values using the comparative C_T (ΔΔCt) method as previously described [14]. The specificity of assay was evaluated by melting curve analysis comprising melting curve ramping from 65 °C to 95 °C with an increasing temperature of 0.5 °C for 5 s (1 cycle).

Results and discussion

Differential biosynthesis and accumulation of swertiamarin, mangiferin and amarogentin

For expression analysis of swertiamarin/amarogentin and mangiferin biosynthesis pathway genes tissue samples from 2-year-old greenhouse plants (roots, leaves and flowers) and 1.5-month old tissue cultured plants (roots and leaves) were used. These tissues varied for the aforementioned secondary metabolites as analyzed by high performance liquid chromatography (HPLC). Swertiamarin, mangiferin and amarogentin content across all tissues ranged from 0.1% to 2.8% of fresh weight (fr. wt), 0.02% to 1.0% of fr. wt and 0.01% to 0.1% of fr. wt, respectively (**Fig. 4**).

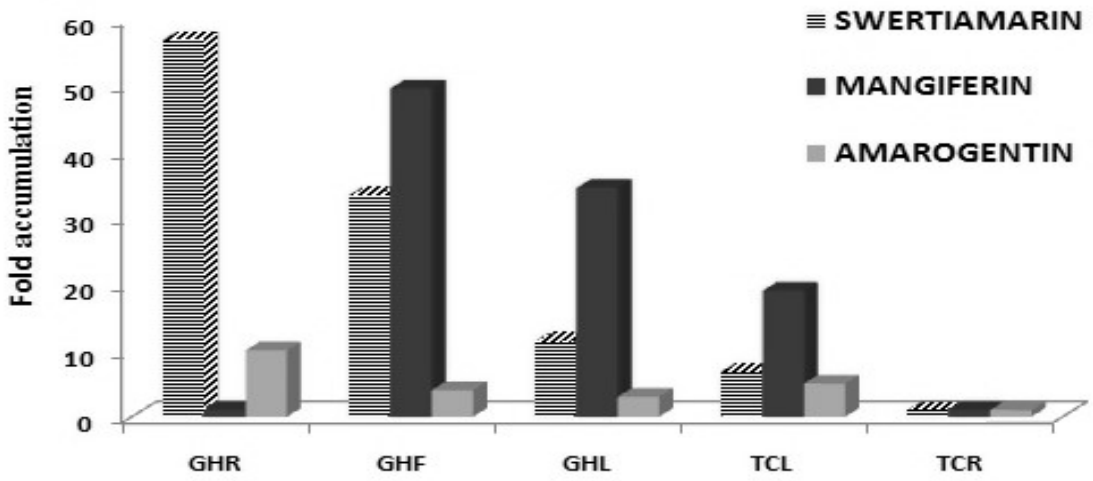

Figure 4. Fold difference in swertiamarin, mangiferin and amarogentin contents in different tissues w.r.t. TCR. GHR, roots of greenhouse grown *Swertia*; GHF, flower of greenhouse grown *Swertia*; GHL, leaves of greenhouse grown *Swertia*; TCL, leaves of tissue cultured *Swertia*; TCR, roots of tissue cultured *Swertia*.

Among all tissues, roots of greenhouse grown plant (GHR) showed the highest content of swertiamarin (2.8% of fr. wt) and amarogentin (0.1% of fr. wt). The leaves of tissue cultured plant (TCL) contained a higher amount of swertiamarin (0.3% of fr. wt) and amarogentin (0.1% of fr. wt) in comparison to its root tissues (TCR). Similarly, maximum mangiferin content (1.0% of fr. wt) was observed in flowers of greenhouse grown plants (GHF). It is inferred from the observation that the biosynthesis and accumulation of secondary metabolites is tissue and organ specific, and dependent on environmental conditions, which was in consistent with previous studies on *Picrorhiza kurroa* [15] and *Artemisia annua* L [16]. All these experiments provided tissues with differential accumulating conditions of the secondary metabolites to be used for expression analysis of biosynthesis pathway genes. For transcriptome analysis composite tissues (shoot and root tissues) of greenhouse grown and *in vitro* cultured Chirata plants were considered. HPLC analysis showed that the composite Chirata tissue contained swertiamarin (3.3% of fr. wt), amarogentin (0.1% of fr. wt) and mangiferin (0.5% of fr. wt), whereas the tissue cultured sample contained 0.4%, 0.02% and 0.2% of the above metabolites, respectively (**Fig. 5**).

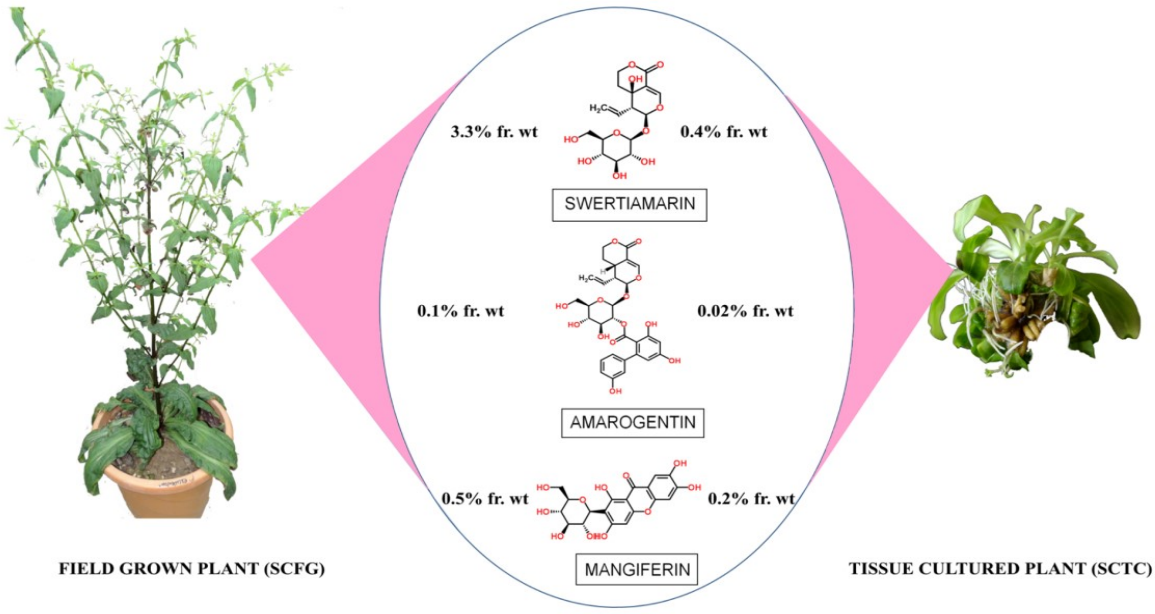

Figure 5. Contents of major secondary metabolites in *Swertia chirayita* plants grown in green-house conditions (photoautotrophic, SCFG) and tissue culture (photoheterotrophic, SCTC)

Expression profiling of pathway genes in relation to metabolite contents

Expression analysis of 24 genes of swertiamarin/amarogentin and 15 genes of mangiferin biosynthesis pathways were carried out in 5 different tissues of *S. chirayita* through RT-qPCR. Initial observation showed the activeness of swertiamarin/amarogentin biosynthesis pathway genes in tissues from greenhouse grown plants as compared to their *in vitro* cultured counterparts. However, this pattern was not observed in mangiferin biosynthesis pathway genes. In roots of greenhouse grown *Swertia* (GHR), which showed the highest amount of swertiamarin and amarogentin, 16 genes of the secoiridoids biosynthesis pathway, encoding DL7H, IS, ISPD, PMK, LMT, MVK, GES, 8HGO, G10H, GDPS, ISPH, 7DLGT, IO, ISPE, IPPI and HMGR showed elevated levels of expression (≥3-fold) in comparison to TCR, with the least content of secoiridoids (**Fig. 6, Table 1**). Among these genes, *DL7H* showed 273 fold, *IS* showed 105 fold, *ISPD* showed 102 fold, *PMK* 57-fold, *LMT* 30-fold, *MVK* 27-fold, *GES* 24-fold, *8HGO* 20-fold, *G10H* 19-fold, *GDPS* 15-fold, *ISPH* 12-fold, *7DLGT* 8-fold, *ISPE* 6-fold and *HMGR* 3-fold increase in transcript abundance. A similar pattern was inferred from leaves of greenhouse grown *Swertia* (GHL), having 11 and 3-fold higher swertiamarin and amarogentin content, respectively, than TCR, which showed 13 genes (encoding MVK, PMK, DXR, ISPD, ISPE, ISPF, IS, LMT, ISPH, IPPI, GDPS, G10H and 8HGO) with more than 3-fold increase in gene expression. Among these genes encoding 8HGO, IS, ISPH, PMK, GDPS, LMT, G10H, IPPI, ISPF, DXR, MVK, ISPD and ISPE showed 35, 22, 16, 15, 13, 10, 9, 8, 7, 7, 5, 4 and 3-fold change in gene expression, respectively. Also, 4 genes (encoding ISPH, 8HGO, DL7H and SLS) showed an increase in gene expression level (≥3-fold) in TCL with respect to TCR. In contrast to these, the floral tissue GHF, showed only 2 genes (coding for ISPH and 8HGO) with more than 3-fold gene expression despite a higher level of swertiamarin and amarogentin (33 and 4-fold, respectively w.r.t TCR). In order to filter out the important genes of the biosynthesis pathway, a comparative study of expression of different genes of the metabolic pathways across different tissues was performed (**Fig. 7**). The analysis showed that the expression of 4 genes (encoding MVK, PMK, GES and LMT) in the swertiamarin/amarogentin biosynthesis varied in accordance with the metabolite concentrations in different tissues.

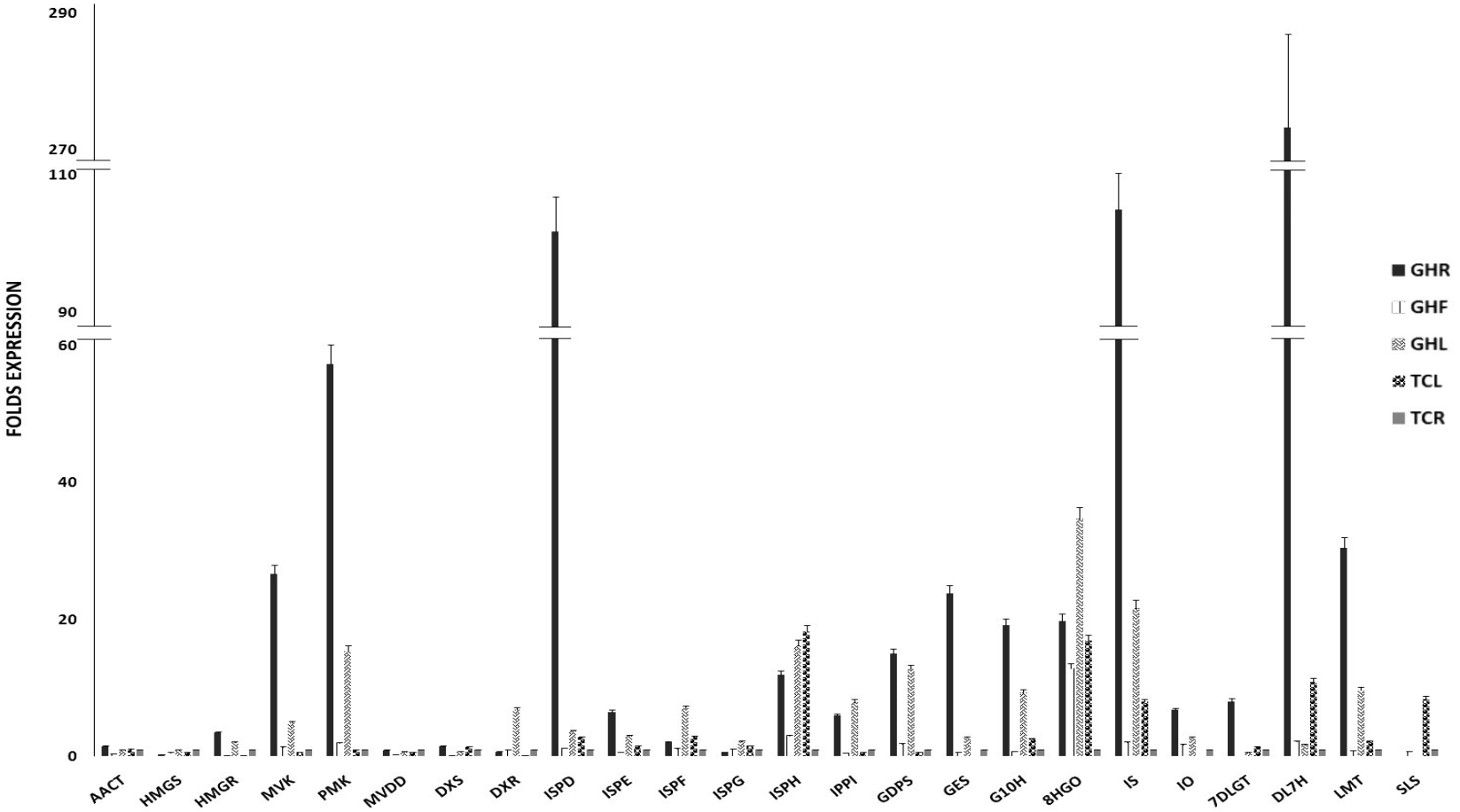

Figure 6. Relative fold changes in the expression of swertiamarin/amarogentin biosynthesis pathway genes across different tissues w.r.t. TCR. GHR, roots of greenhouse grown *Swertia*; GHF, flower of greenhouse grown *Swertia*; GHL, leaves of greenhouse grown *Swertia*; TCL, leaves of tissue cultured *Swertia*; TCR, roots of tissue cultured *Swertia*

Table 1. Relative fold expression of secoiridoid biosynthesis pathway genes across different tissues with respect to roots of tissue cultured *Swertia* (TCR).

		Greenhouse Roots (GHR)	Greenhouse Flowers (GHF)	Greenhouse Leaves (GHL)	Tissue cultured Leaves (TCL)
1	*AACT*	-	-	-	-
2	*HMGS*	-	-	-	-
3	*HMGR*	+	-	-	-
4	*MVK*	++	-	+	-
5	*PMK*	+++	-	++	-
6	*MVDD*	-	-	-	-
7	*DXS*	-	-	-	-
8	*DXR*	-	-	+	-
9	*ISPD*	++++	-	+	-
10	*ISPE*	+	-	+	-
11	*ISPF*	-	-	+	-
12	*ISPG*	-	-	-	-
13	*ISPH*	++	+	++	++
14	*IPPI*	+	-	+	-
15	*GDPS*	++	-	++	-
16	*GES*	++	-	-	-
17	*G10H*	++	-	+	-
18	*8HGO*	++	++	++	++
19	*IS*	++++	-	++	+
20	*IO*	+	-	-	-
21	*7DLGT*	+	-	-	-
22	*DL7H*	++++	-	-	++
23	*LMT*	++	-	+	-
24	*SLS*	-	-	-	+
Scale:	<3 fold (-) ≥3-10 fold (+) ≥10-40 fold (++) ≥40-80 (+++) ≥80 (++++)				

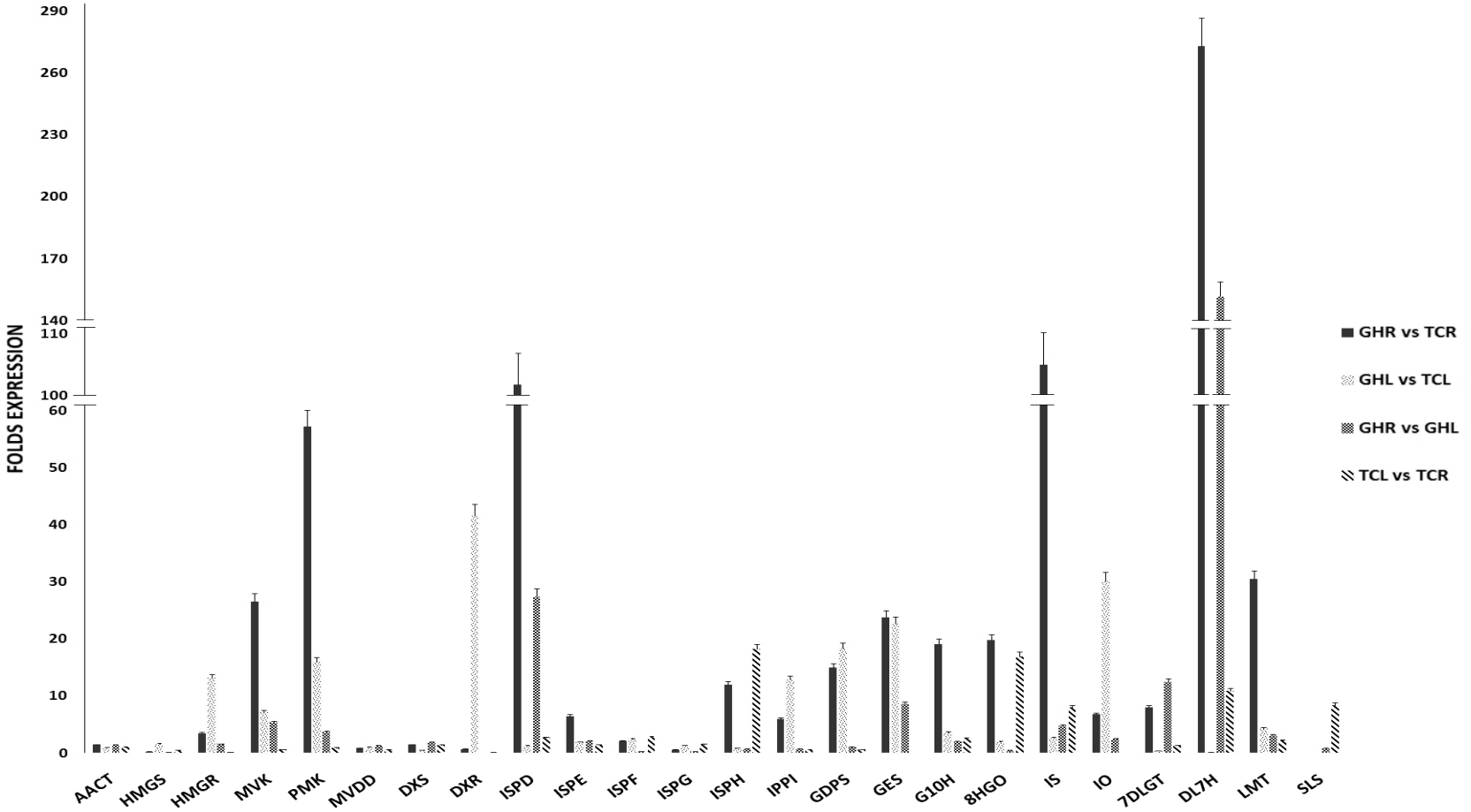

Figure 7. A comparative account of expression of pathway genes across different tissues. GHR, roots of greenhouse grown *Swertia*; GHF, flower of greenhouse grown *Swertia*; GHL, leaves of greenhouse grown *Swertia*; TCL, leaves of tissue cultured *Swertia*; TCR, roots of tissue cultured *Swertia*.

The expression analysis of 15 genes participating in the mangiferin biosynthesis revealed some correlation of a number of genes in tune with the metabolite contents and the tissue types (**Fig. 8** and **Table 2**). In GHR, 10 genes of the metabolic pathway (coding for DAHPS, DHQS, DHQD, EPSPS, CS, CM, ADT, PAL, C4H and AAC) showed increased levels (≥3-fold) of expression in defiance of the low content of mangiferin. The transcript abundance of 8 genes (encoding EPSPS, DHQS, PAL, ADT, CM, DAHPS, DHQD and CS) increased in relation to mangiferin content in leaves. However, in floral tissue, transcript abundance of most of the genes except *EPSPS*, dropped in spite of mangiferin content being the highest, thereby suggesting that the floral tissue may be a site of mangiferin accumulation rather biosynthesis. A subsequent PCA analysis (data not shown) unveiled 10 genes of secoiridoid biosynthesis (encoding HMGR, GES, 7DLGT, 8HGO, ISPD, ISPE, PMK, MVK, IS and G10H) and 5 genes (coding for EPSPS, CS, PAL, CM and ADT) of mangiferin biosynthesis as the major contributors. The overall study identified certain prime gene candidates encoding HMGR, PMK, MVK, ISPD, ISPE, GES, G10H, 8HGO, IS and 7DLGT of the secoiridoids biosynthesis pathway and 5 genes (coding for EPSPS, PAL, ADT, CM and CS) of mangiferin biosynthesis significantly contributing to the biosynthesis of these metabolites.

Our study was analogous to previous reports, where secondary metabolites like picrosides in *P. kurroa* [15], shikonin in *Arnebia euchroma* [17] and aconites in *Aconitum heterophyllum* [18] have been shown to be positively correlated with the expression profile of multiple genes of their respective metabolic pathways. Higher expression levels of *ISPD* and *ISPE* have been associated with higher picroside-I content in *P. kurroa* [15]. Similar expression studies revealed the role of HMGR in *A. annua* [19], where overexpression of the gene enhanced the production of artemisinin. The importance of genes coding for PMK and MVK has been well illustrated in previous studies, where they were targeted to enhance the overall isoprenoid production [20]. Overexpression of *G10H* in the hairy roots has improved production of several monoterpene indole alkaloids (MIAs) [21]. The significance of gene encoding GES in geraniol production has been described in a previous report, where they optimized geraniol production based on expression of plastid-targeted *GES* from *Valeriana officinalis* in transgenic tobacco suspension cultures [22]. Apart from the leaves, expression profiling of genes involved in mangiferin biosynthesis did not show any specific pattern in relation to mangiferin content across the

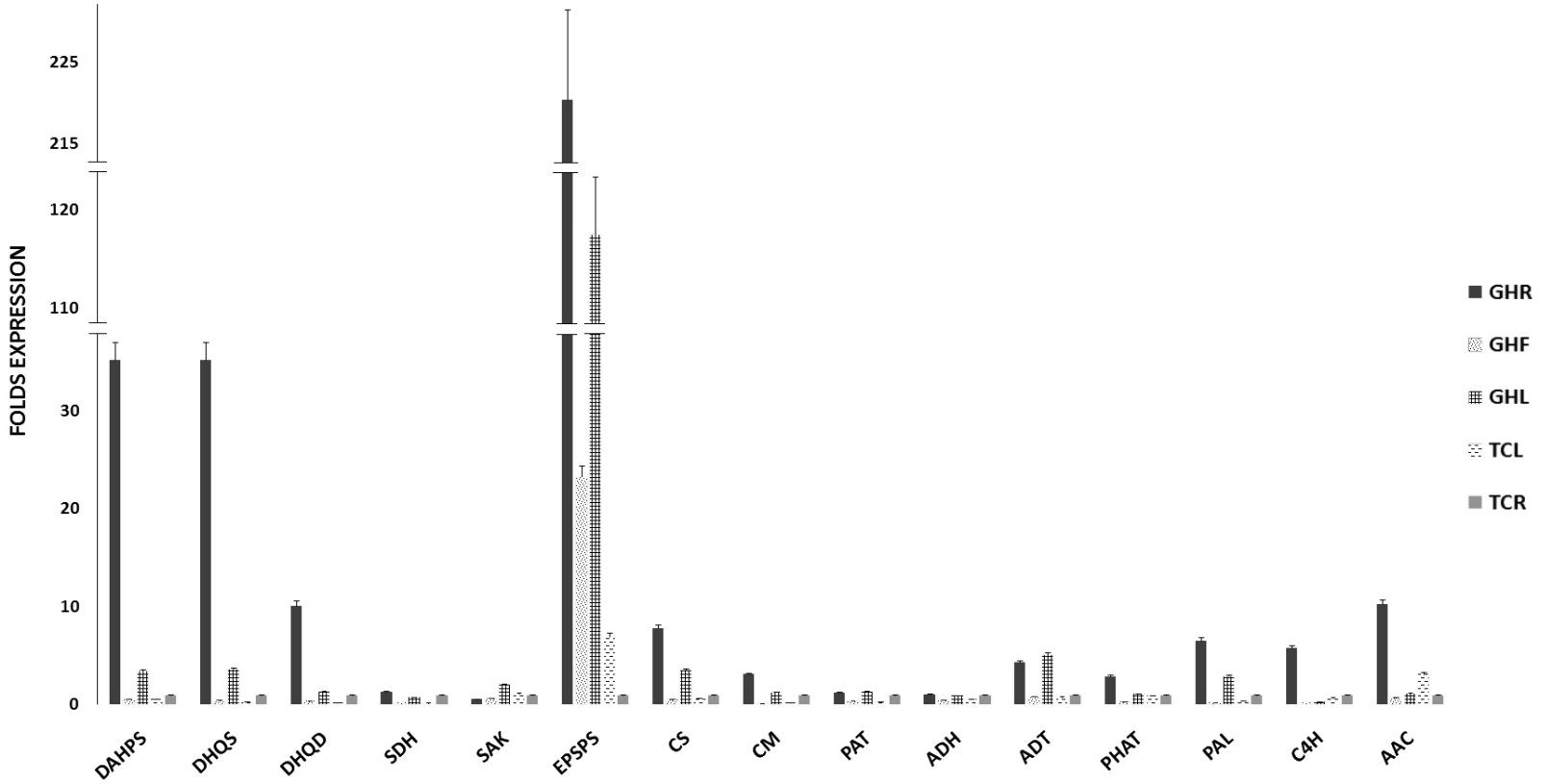

Figure 8. Relative fold changes in the expression of mangiferin biosynthesis pathway genes across different tissues w.r.t. TCR

Table 2. Relative fold expression of mangiferin biosynthesis pathway genes across different tissues with respect to roots of tissue cultured *Swertia* (TCR).

		Greenhouse Roots (GHR)	Greenhouse Flowers (GHF)	Greenhouse Leaves (GHL)	Tissue cultured Leaves (TCL)
1	*DAHPS*	++	-	+	-
2	*DHQS*	++	-	+	-
3	*DHQD*	++	-	-	-
4	*SDH*	-	-	-	-
5	*SAK*	-	-	-	-
6	*EPSPS*	++++	++	++++	+
7	*CS*	+	-	+	-
8	*CM*	+	-	-	-
9	*PAT*	-	-	-	-
10	*ADH*	-	-	-	-
11	*ADT*	+	-	+	-
12	*PHAT*	-	-	-	-
13	*PAL*	+	-	-	-
14	*C4H*	+	-	-	-
15	*AAC*	++	-	-	+
Scale: <3 fold (-) ≥3-10 fold (+) ≥10-40 fold (++) ≥40-80 (+++) ≥80 (++++)					

investigated tissues. This is in accordance with a previous report, where the expression levels of the genes involved in the general phenylpropanoid pathway did not show significant correlation with the accumulation pattern of shikonin [23]. This may suggest that the highly expressed genes are involved in the biosynthesis of other secondary metabolites. In consistent with our finding, a previous study also indicated the importance of *PAL* gene, whose expression varied in accordance with condensed tannins (CT) and lignin biosynthesis in developing shoot and root tissues of *Populus tremuloides* Michx. [24]. The *EPSPS* gene, which emerged as a potential regulatory gene in our study, has been reported to be a significant player in the aromatic amino acid biosynthesis pathway being used as glyphosate target [25]. The floral tissue did not show correlation of the expression pattern of pathway genes in relation to their metabolite contents. This may suggest the involvement of a transport mechanism, wherein the metabolites are

synthesized outside the sink-tissue and are later transported to floral tissue, in a similar manner like nicotine gets transported from roots to the central vacuoles of leaves by Nt-JAT1, a (MATE)-type transporter in *Nicotiana tabacum* [26], and transport of glucosinolates from maternal tissue to seeds involving NRT/PTR transporters of *Arabidopsis thaliana* [27].

Fundamental processes like glycolysis, tricarboxylic acid (TCA) cycle, pentose phosphate pathway, photosynthesis etc., are involved in the biosynthesis, degradation and interconversion of primary and secondary metabolites [28]. Majority of building blocks of secondary metabolic pathways (acetate, shikimate, mevalonate and deoxyxylulose pathways etc.) are derived from important intermediates such as acetyl Co-A, shikimic acid, mevalonic acid, 1-deoxyxylulose 5-phosphate, glucose 6-phosphate, glyceraldehyde 3-phosphate, phosphoenolpyruvate and erythrose 4-phosphate. Therefore, it is logical to assume that the formation of these metabolites is probably influenced by the state and efficiencies of these fundamental processes which provide precursors to secondary metabolites biosynthesis. Lower levels of secondary metabolites observed in photoheterotrophic cultures of Chirata may be explained based on their lesser photosynthetic capacity as compared to the photoautotrophic plants.

For expression analysis, sixteen genes (coding for IDH, G6PD, PRK, OGDH, DLD, DLST, HXK, PFK1, PGK, FBP, RUBISCO, ENO, FUM, GPI, SUCLG and RPE), involved in various fundamental processes were selected on the basis of higher *in silico* transcript abundance values. RT-qPCR analysis revealed 10 genes (encoding G6PD, OGDH, DLD, HXK, PFK1, FBP, RUBISCO, FUM, SUCLG and RPE) out of 16 showed elevated levels of expression ($\geq$ 2 fold) in SCFG in comparison to SCTC (**Fig. 9**), which was found to be largely in agreement with the RNA-seq data. The differentially regulated multiple pathways encountered in differential growth regimes (SCFG and SCTC) are suggested to be contributing towards growth and secondary metabolites biosynthesis in *S. chirayita.*

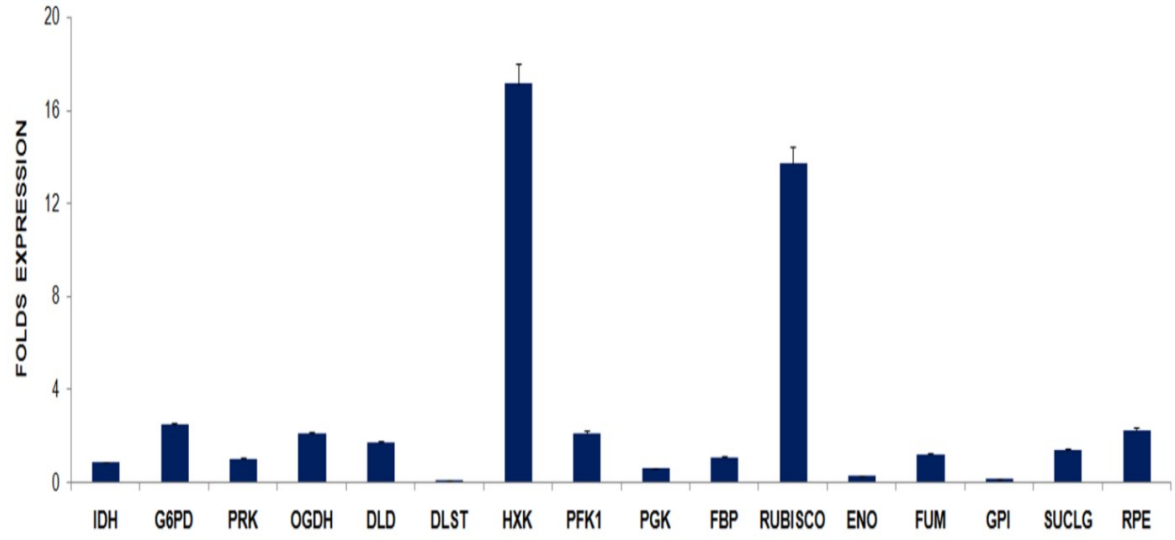

Figure 9. Relative fold changes in the expression of fundamental pathway genes in SCFG w.r.t SCTC. Error bars correspond to percentage error (5 %).

There are growing apprehensions over the demand-supply of *Swertia*-based raw materials that is mostly affected by conventional harvesting practices and less-successful *in vitro* techniques. Poorly understood biosynthetic pathways of these secondary metabolites and the regulatory networks thereof, add more concern to the whole scenario. Of the various means the gene gets regulated, small RNA (sRNA) mediated regulatory surveillance has occupied a major share of research areas displaying their potential in growth, development and various defense mechanisms. In this report, we have computationally mined *S. chirayita* transcriptomes generated from tissues with contrasting levels of major medicinal metabolites, and identified putative miRNAs having possible association with the biosynthesis of secondary metabolites in Chirata. A schematic representation of the methodologies followed in identification and experimental validation of miRNA and their targets is given in **Fig. 10**.

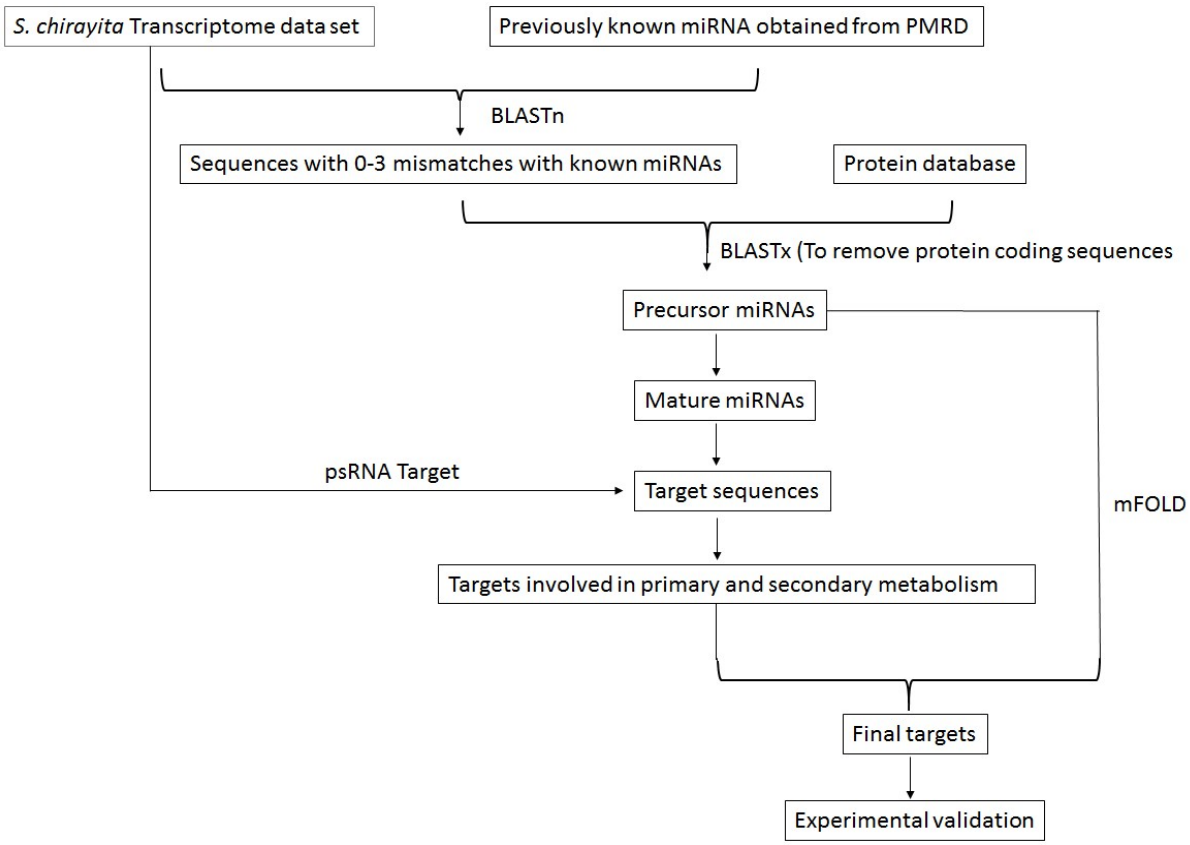

Figure 10. Workflow depicting the methodologies followed in the computational identification of miRNAs and their target prediction in *S. chirayita.*

A total of ten miRNAs and eleven potential targets were chosen based on the default filtration criterion set using mfold server. Out of the ten miRNAs; *miR-168, miR-11320* and *miR-166a* are the miRNAs found to be directly associated with secondary metabolism targeting *acetyl-CoA acetyltransferase* (*AACT*), *aspartate aminotransferase* (*PHAT*) and *premnaspirodiene oxygenase* (*PSO*). *miR-166b* is another miRNA identified in our study that targeted a gene encoding homeobox-leucine zipper protein (*HD-ZIP*), a transcription factor family, which plays a significant role in secondary metabolism. Other miRNAs, such as *miR-156a, miR-11071, miR-162, miR-171* and *miR-396* have targets mostly of the primary metabolic processes which provide building blocks for the secondary metabolism. Targets of the primary metabolic processes are of utmost importance as they provide precursors for various secondary metabolism [28]. A schematic representation of the miRNAs and their possible targets has been illustrated in **Fig.11**.

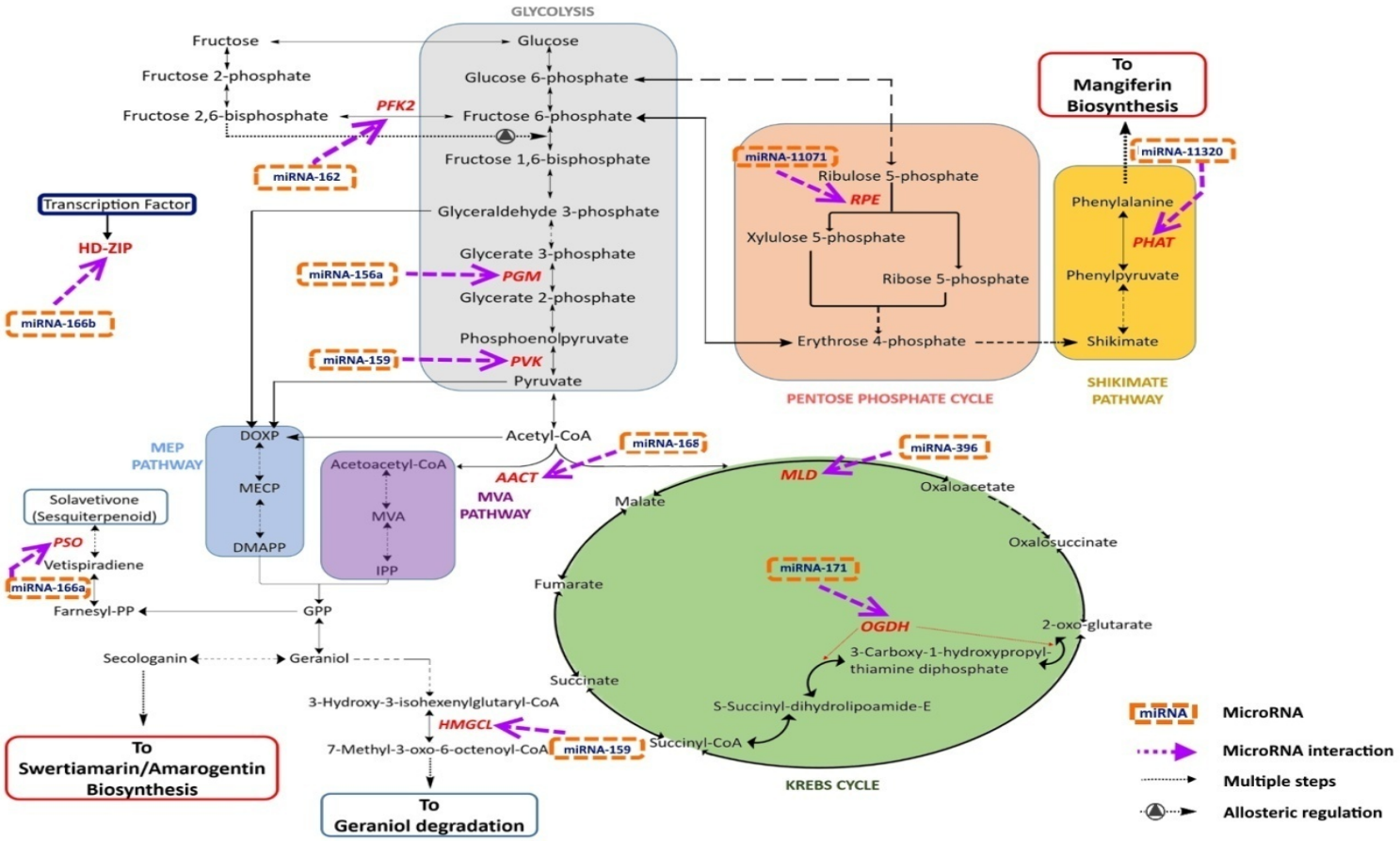

Figure 11. Schematic representation of the possible roles played by the putative miRNAs and their targets in secondary metabolism in *Swertia chirayita*

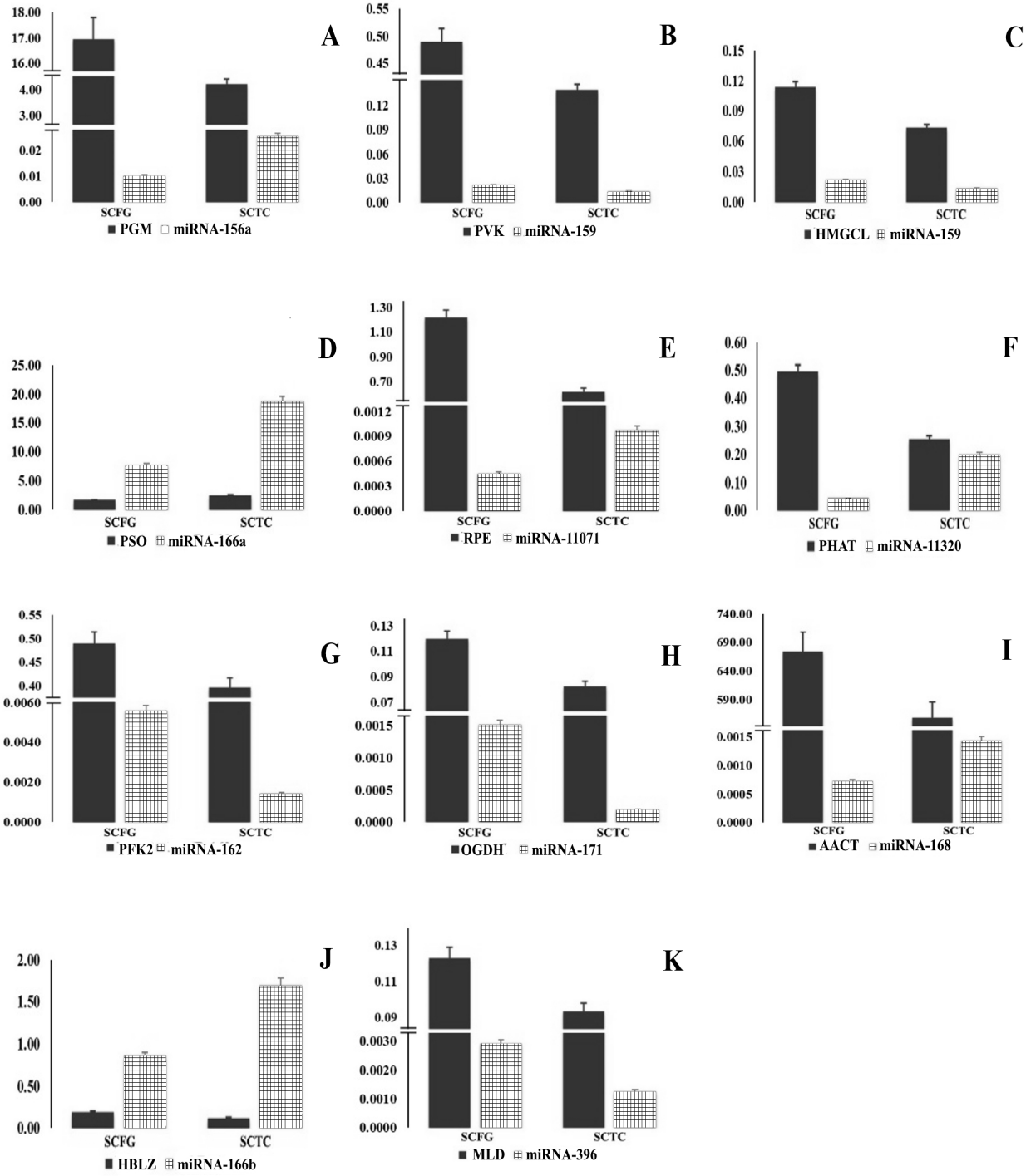

Figure 12. Expression status of miRNAs and their targets in greenhouse grown (SCFG) and *in vitro* cultured (SCTC) *S. chirayita* tissues

Tissue specific roles of putative miRNAs and their targets were analyzed through RT-qPCR using specific primers. The results showed expression of all the miRNAs in tissues of greenhouse grown and tissue cultured plants (**Fig. 12**). Six miRNAs; *miR-168, miR-11320,*

miR-166a, miR-11071, miR-156a and *miR-166b*, targeting genes encoding acetyl-CoA acetyltransferase (AACT), aspartate aminotransferase (PHAT), premnaspirodiene oxygenase (PSO), ribulose-phosphate 3-epimerase (RPE), phosphoglycerate mutase (PGM) and a homeobox-leucine zipper protein (HD-ZIP), respectively showed higher expression in tissue cultured plants. On the other hand, microRNAs *miR-159, miR-162, miR-396* and *miR-171* having gene targets coding for pyruvate kinase (PVK) and hydroxymethylglutaryl-CoA lyase (HMGCL), phosphofructo kinase 2 (PFK2) and malate dehydrogenase (MLD) and 2-oxoglutarate dehydrogenase E1 component (OGDH), respectively were present in higher amount in greenhouse grown plant tissues. Maximum expression of *miR-166a* followed by *miR-166b* was observed in tissue cultured plants and least expression level was exhibited by *miR-11071* in the same tissue. On the other hand, almost all targets showed a higher fold in expression in tissues procured from greenhouse grown plants, which can be attributed and correlated to the presence of higher levels of secondary metabolites. The computationally identified putative miRNAs could serve as suitable target elements aiming at enhancing secondary metabolites level in endangered medicinal plants like *S. chirayita.*

Concluding remark and future prospects

The study opened up new windows towards understanding the molecular basis of biosynthesis of secondary metabolites in *S. chirayita*. This was the first report wherein the expression profiling of swertiamarin/amarogentin and mangiferin biosynthesis pathway genes were analyzed in relation to metabolites contents in *S. chirayita*. Potential key genes with elevated transcript abundance in relation to metabolite contents could be used as suitable targets for the development of gene markers for genetic improvement of *S. chirayita* with enhanced secondary metabolite production. Through the study of greenhouse grown (photoautotrophic) versus the tissue cultured (photoheterotrophic) Chirata plants, we have made an attempt to link the contribution of important, yet much less explored fundamental processes to secondary metabolites biosynthesis in plant systems. Optimization of factors (intrinsic and extrinsic), those decide the final metabolite levels, and devising suitable molecular strategies would help *in vitro* studies in *S. chirayita* for enhancing secondary metabolites level. miRNA mediated regulation of secondary metabolites biosynthesis pathway genes, after their functional validation, would be useful in planning genetic improvement strategies for enhancing secondary metabolites content in *S. chirayita*.

References

[1] K. Patil, S. Dhande, and V. Kadam, "Therapeutic *Swertia chirata* - An Overview", *Research Journal of Pharmacognosy and Phytochemistry,* vol. 5, pp. 199-207, 2013.

[2] S. Phoboo, M.D.S. Pinto, A.C.L. Barbosa, D. Sarkar, P.C. Bhowmik, P.K. Jha, and K. Shetty, "Phenolic-Linked Biochemical Rationale for the Anti-Diabetic Properties of *Swertia chirayita* (Roxb. ex Flem.) Karst.", *Phytotherapy Research,* vol. 27, pp. 227-235, 2013.

[3] K. Miettinen, L. Dong, N. Navrot, T. Schneider, V. Burlat, J. Pollier, L. Woittiez, S. van der Krol, R. Lugan, and T. Ilc, R. Verpoorte, K.M. Oksman-Caldentey, E. Martinoia, H. Bouwmeester, A. Goossens, J. Memelink, and D. Werck-Reichhart "The seco-iridoid pathway from *Catharanthus roseus*", *Nature communications,* vol. 5, 2014.

[4] C.Z. Wang, U.H. Maier, W. Eisenreich, P. Adam, I. Obersteiner, M. Keil, A. Bacher, and M.H. Zenk, "Unexpected Biosynthetic Precursors of Amarogentin − A Retrobiosynthetic ^{13}C NMR Study", *European Journal of Organic Chemistry,* vol. 2001, pp. 1459-1465, 2001.

[5] H. Kuwajima, N. Hayashi, K. Takaishi, K. Inoue, Y. Takeda, and H. Inouye, "Studies on monoterpene glucosides and related compounds. LXV. Biosynthesis of the biphenylcarboxylic acid moiety of amarogentin and amaroswerin", *Yakugaku Zasshi= Journal of the Pharmaceutical Society of Japan,* vol. 110, pp. 484-489, 1990.

[6] M. Fujita and T. Inoue, "Biosynthesis of mangiferin in *anemarrhena asphodeloides*: intact incorporation of C_6-C_3 precursor into xanthone", *Tetrahedron Letters,* vol. 18, pp. 4503-4506, 1977/01/01 1977.

[7] S. Karuppusamy, "A review on trends in production of secondary metabolites from higher plants by in vitro tissue, organ and cell cultures", *Journal of Medicinal Plants Research*, vol. 3(13), pp. 1222-1239, 2009.

[8] T. Murashige and F. Skoog, "A revised medium for rapid growth and bio assays with tobacco tissue cultures", *Physiologia plantarum,* vol. 15, pp. 473-497, 1962.

[9] V. Kumar, R. Chauhan, and H. Sood, "In vitro production and efficient quantification of major phytopharmaceuticals in an endangered medicinal herb, *Swertia chirata*", *Int J Biotechnol Bioeng Res,* vol. 4, pp. 495-506, 2013.

[10] A. Das and T.K. Mondal, "Computational identification of conserved microRNAs and their targets in tea (*Camellia sinensis*)", *American Journal of Plant Sciences,* vol. 1, p. 77, 2010.

[11] F. Xie, T.P. Frazier, and B. Zhang, "Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*)", *Planta,* vol. 232, pp. 417-434, 2010.

[12] S. Rozen and H. Skaletsky, "Primer3 on the WWW for general users and for biologist programmers", *Methods in molecular biology*, vol. 132, 365-386, 2000.

[13] M.W. Pfaffl, A. Tichopad, C. Prgomet, and T.P. Neuvians, "Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper - Excel-based tool using pair-wise correlations", *Biotechnology letters,* vol. 26, pp. 509-515, 2004.

[14] T.D. Schmittgen and K.J. Livak, "Analyzing real-time PCR data by the comparative C_T method", *Nature protocols,* vol. 3, pp. 1101-1108, 2008.

[15] S. Pandit, K. Shitiz, H. Sood, P.K. Naik, and R.S. Chauhan, "Expression pattern of fifteen genes of non-mevalonate (MEP) and mevalonate (MVA) pathways in different tissues of endangered medicinal herb *Picrorhiza kurroa* with respect to picrosides content," *Molecular biology reports",* vol. 40, pp. 1053-1063, 2013.

[16] L. Olofsson, A. Engström, A. Lundgren, and P.E. Brodelius, "Relative expression of genes of terpene metabolism in different tissues of *Artemisia annua* L", *BMC plant biology,* vol. 11, p. 1, 2011.

[17] R.S. Singh, R.K. Gara, P.K. Bhardwaj, A. Kaachra, S. Malik, R. Kumar, M. Sharma, P.S. Ahuja, and S. Kumar, "Expression of *3-hydroxy-3-methylglutaryl-CoA reductase*, *p-hydroxybenzoate-m-geranyltransferase* and genes of phenylpropanoid pathway exhibits positive correlation with shikonins content in arnebia [*Arnebia euchroma* (Royle) Johnston]", *BMC molecular biology,* vol. 11, p. 1, 2010.

[18] N. Malhotra, V. Kumar, H. Sood, T.R. Singh, and R.S. Chauhan, "Multiple genes of mevalonate and non-mevalonate pathways contribute to high aconites content in an endangered medicinal herb, *Aconitum heterophyllum* Wall", *Phytochemistry,* vol. 108, pp. 26-34, 2014.

[19] P. Alam and M. Abdin, "Over-expression of HMG-CoA reductase and amorpha-4, 11-diene synthase genes in *Artemisia annua* L. and its influence on artemisinin content", *Plant cell reports,* vol. 30, pp. 1919-1928, 2011.

[20] P. Singh, T.S. Batth, D. Juminaga, R.H. Dahl, J.D. Keasling, P.D. Adams, and C.J. Petzold, "Application of targeted proteomics to metabolically engineered *Escherichia coli*", *Proteomics,* vol. 12, pp. 1289-1299, 2012.

[21] C.A.M. Peebles, G.W. Sander, E.H. Hughes, R. Peacock, J.V. Shanks, and K.Y. San, "The expression of 1-deoxy-D-xylulose synthase and geraniol-10-hydroxylase or anthranilate synthase increases terpenoid indole alkaloid accumulation in *Catharanthus roseus* hairy roots", *Metabolic engineering,* vol. 13, pp. 234-240, 2011.

[22] N. Vasilev, C. Schmitz, U. Grömping, R. Fischer, and S. Schillberg, "Assessment of cultivation factors that affect biomass and geraniol production in transgenic tobacco cell suspension cultures", *PloS one,* vol. 9, p. e104620, 2014.

[23] Y. Yamamura, F.P. Sahin, A. Nagatsu, and H. Mizukami, "Molecular cloning and characterization of a cDNA encoding a novel apoplastic protein preferentially expressed in a shikonin-producing callus strain of *Lithospermum erythrorhizon*", *Plant and cell physiology,* vol. 44, pp. 437-446, 2003.

[24] Y.Y. Kao, S.A. Harding, and C.J. Tsai, "Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen", *Plant physiology,* vol. 130, pp. 796-807, 2002.

[25] Y.S. Tian, J. Xu, J. Han, W. Zhao, X.Y. Fu, R.H. Peng, and Q.H. Yao, "Complementary screening, identification and application of a novel class II 5-enopyruvylshikimate-3-phosphate synthase from *Bacillus cereus*", *World Journal of Microbiology and Biotechnology,* vol. 29, pp. 549-557, 2013.

[26] N. Shitan, M. Morita, and K. Yazaki, "Identification of a nicotine transporter in leaf vacuoles of *Nicotiana tabacum*", *Plant signaling & behavior,* vol. 4, pp. 530-532, 2009.

[27] H.H. Nour-Eldin, T.G. Andersen, M. Burow, S.R. Madsen, M.E. Jørgensen, C.E. Olsen, I. Dreyer, R. Hedrich, D. Geiger, and B.A. Halkier, "NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds", *Nature,* vol. 488, pp. 531-534, 2012.

[28] P.M. Dewick, "Secondary metabolism: the building blocks and construction mechanisms", *Medicinal Natural Products: A Biosynthetic Approach, Second Edition,* pp. 7-34, 2001.

List of publications

Journal Publications

1. **J.K. Padhan**, V. Kumar, H. Sood, T.R. Singh, and R.S. Chauhan (2015). Contents of therapeutic metabolites in *Swertia chirayita* correlate with the expression profiles of multiple genes in corresponding biosynthesis pathways. *Phytochemistry*, 116:38-47 doi: 10.1016/j.phytochem.2015.05.007 **(IF - 2.547, Scopus Indexed)**

2. **J.K. Padhan,** P. Kumar, H. Sood and R.S. Chauhan (2016). Prospecting NGS-transcriptomes to assess regulation of miRNA-mediated secondary metabolites biosynthesis in *Swertia chirayita*, a medicinal herb of North-Western Himalayas. *Medicinal Plants - International Journal of Phytomedicines and Related Industries*, 8(3):219-228 doi: 10.5958/0975-6892.2016.00029.0 (IF – N/A, **Scopus Indexed**)

3. **J.K. Padhan**, T. Pal, P. Kumar, H. Sood and R.S. Chauhan (2016). Comparative transcriptomics uncovers differences in photoautotrophic versus photoheterotrophic modes of nutrition in relation to secondary metabolites biosynthesis in *Swertia chirayita*. *Molecular Biology Reports,* (**IF - 1.698, Scopus Indexed**) (Under review)

Conference Publications

1. **J.K. Padhan,** V. Kumar, H. Sood and R.S. Chauhan (2014). Expression profiling of swertiamarin and amarogentin biosynthetic pathway genes in an endangered medicinal herb, *Swertia chirayita* (Roxb. Ex Fleming) H. Karst. Proceedings of the National Symposium on Advances in Biotechnology for Crop Improvement [Baru Sahib, HP : 12 July, 2014].

2. **J.K. Padhan,** V. Kumar, H. Sood and R.S. Chauhan (2014). Mangiferin biosynthesis in *Swertia chirayita* is correlated to elevated expression of phenylpropanoid pathway genes. Proceedings of the International Conference on Emerging Trends in Biotechnology [New Delhi, India: 6-9 November 2014].

3. **J.K. Padhan** and R.S. Chauhan (2015). Transcription factor mediated regulatory surveillance of therapeutic metabolite biosynthesis in *Swertia chirayita*, an endangered medicinal herb of North-Western Himalayas. Proceedings of the National Symposium on Germplasm to Genes: Harnessing Biotechnology for Food Security and Health [National Agricultural Science Centre, Pusa, New Delhi, India: August 9-11, 2015].